CAPSTONE PROJECT: IBM DATA SCIENCES PROFESSIONAL CERTIFICATE.

THE BATTLE OF NEIGHBORHOODS: CHICAGO.

May, 2020

By: Rómulo Guédez

THE BATTLE OF NEIGHBORHOODS: CHICAGO.

1. INTRODUCTION.
2. BUSINESS PROBLEM.
3. DATA SECTION.
4. METHODOLOGY.
5. RESULTS.
6. DISCUSSION.
7. CONCLUSIONS.

1. INTRODUCTION

CHICAGO IS THE THIRD LARGEST CITY IN THE UNITED STATES OF AMERICA, WITH A POPULATION OF NEARLY THREE MILLION PEOPLE, AND IS HOME TO SEVENTY-SEVEN (77) COMMUNITY AREAS AND ONE HUNDRED (100) NEIGHBORHOODS https://www.chicago.gov/city/en/about/facts.html.

ONE OF ITS MOST IMPORTANT FACTS IS THAT THE CITY IS SEING AS A FOODIE DESTINATION, WITH 7 AAA DIAMOND-RATED RESTAURANTS, 26 MICHELIN-STARRED RESTAURANTS, HOST OF THE ANNUAL JAMES BEARD AWARDS AND IT EVEN HAS 144 DOG-FRIENDLY RESTAURANTS.

CHICAGO SKYLINE
(https://sports.yahoo.com/chicago-travel-review-art-architecture-165852857.html)

MILLENIUM PARK
(https://www.titanpainterschicago.com/blog/best-places-to-go-in-chicago)

2. BUSINESS PROBLEM.

- AN IMPORTANT ENTREPRENEUR BASED ON THE EAST COAST, WHICH ITS MAIN SPECIALIZATION IS THE BRANCH OF RESTAURANTS, IS INTERESTED TO EXPAND THEIR BUSINESS TO OTHER REGIONS OF THE COUNTRY.
- THE MAIN AIM OF THIS PROJECT IS PROVIDE SOME RECOMMENDATIONS TO THE CLIENT, IN ORDER TO HELP THEM TO RESOLVE THE PROBLEM OF OPENING NEW RESTAURANTS IN THE CITY OF CHICAGO.
- THE RECOMMENDATIONS WILL BE BASE IN THIS FEASIBILITY STUDY RESULTS, WHICH CONSIST TO ANALYZE THE CURRENT DISTRIBUTION OF RESTAURANTS CATEGORIES ALONG THE CITY. THIS COULD LED US TO VISUALIZE SOME HIDDEN PATTERNS, SUCH AS A POSSIBLE SEGMENTATION OF THE RESTAURANT CATEGORIES DISTRIBUTION THROUGH THE CITY. THIS WILL BE A VERY HELPFUL TOOL IN ORDER TO UNDERSTAND THE CURRENT MARKET AND HELP TO MAKE DECISIONS ABOUT, FOR INSTANCE, PROPOSE PRELIMINARY LOCATIONS FOR A DETERMINED KIND OF RESTAURANT IN A SPECIFIC PLACE.

3. DATA SECTION

- Data 1: Neighborhoods Information.
  - Neighborhoods information is required to learn about the city´s urban organization and segment its populate areas. Chicago is home of 77 community areas and more than 100 Neighborhoods.
  - the source of Chicago neighborhoods data is the follow internet address:
    https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Chicago.
  - The neighborhood data was organized in a dataframe.

	Neighborhood	Community area
0	Albany Park\n	Albany Park\n
1	Altgeld Gardens\n	Riverdale\n
2	Andersonville\n	Edgewater\n
3	Archer Heights\n	Archer Heights\n
4	Armour Square\n	Armour Square\n

Neighborhood List Dataframe.

3. DATA SECTION

- Data 2: Neighborhoods Coordinates (Longitude & Latitude).
  - Get Longitude and Latitude of neighborhoods using the tool **Nominatin**.

	Neighborhood	Community area	Latitude	Longitude
0	Albany Park\n	Albany Park\n	41.9719	-87.7162
1	Altgeld Gardens\n	Riverdale\n	41.6549	-87.6004
2	Andersonville\n	Edgewater\n	41.9771	-87.6693
3	Archer Heights\n	Archer Heights\n	41.8114	-87.7262
4	Armour Square\n	Armour Square\n	41.84	-87.6331
5	Ashburn\n	Ashburn\n	41.7494	-87.7135
6	Ashburn Estates\n	Ashburn\n	Not Found	Not Found
7	Auburn Gresham\n	Auburn Gresham\n	41.7508	-87.6629
8	Avalon Park\n	Avalon Park\n	41.745	-87.5887
9	Avondale\n	Avondale\n	41.9389	-87.7112

Latitude/Longitude NOT FOUND.

Neighborhoods Name and Latitude/Longitude List Dataframe.

3. DATA SECTION

- Data 2: Neighborhoods Coordinates (Longitude & Latitude).
  - The dataframe was cleaned, deleting those rows without Latitude/Longitude information, and getting a dataframe of 190 rows.

	Neighborhood	Community area	Latitude	Longitude
0	Albany Park\n	Albany Park\n	41.9719	-87.7162
1	Altgeld Gardens\n	Riverdale\n	41.6549	-87.6004
2	Andersonville\n	Edgewater\n	41.9771	-87.6693
3	Archer Heights\n	Archer Heights\n	41.8114	-87.7262
4	Armour Square\n	Armour Square\n	41.84	-87.6331
5	Ashburn\n	Ashburn\n	41.7494	-87.7135
7	Auburn Gresham\n	Auburn Gresham\n	41.7508	-87.6629
8	Avalon Park\n	Avalon Park\n	41.745	-87.5887
9	Avondale\n	Avondale\n	41.9389	-87.7112
11	Back of the Yards\n	New City\n	41.8075	-87.6662
12	Belmont Central\n	Belmont Cragin\n	41.9317	-87.7687
13	Belmont Gardens\n	Hermosa\n	41.9386	-87.7287
15	Belmont Terrace\n	Dunning\n	41.9376	-87.8338
16	Beverly\n	Beverly\n	41.7182	-87.6718
18	Beverly Woods\n	Morgan Park\n	41.6834	-87.6812

Dataframe.shape

```
Out[15]: (190, 4)
```

Neighborhoods Name and Latitude/Longitude List Dataframe Cleaned.

3. DATA SECTION

- Data 3: Venues Information.
  - We used the integrate dataframe of Chicago neighborhoods & geographical coordinates as input for the Foursquare API (www.foursquare.com). Through this API we were able to find the basic venues information for each neighborhood.
  - The basic venues information from Foursquare consist of Venue Name,

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Albany Park\n	41.971937	-87.716174	Cairo Nights Hookah Lounge	41.975776	-87.715547	Hookah Bar
1	Albany Park\n	41.971937	-87.716174	Nighthawk	41.967974	-87.713415	Cocktail Bar
2	Albany Park\n	41.971937	-87.716174	Chicago Produce	41.970553	-87.716327	Grocery Store
3	Albany Park\n	41.971937	-87.716174	Peking Mandarin Resturant	41.968292	-87.715783	Chinese Restaurant
4	Albany Park\n	41.971937	-87.716174	Markellos Baking Company	41.968602	-87.716607	Bakery
5	Albany Park\n	41.971937	-87.716174	Popeyes Louisiana Kitchen	41.968459	-87.713156	Fried Chicken Joint
6	Albany Park\n	41.971937	-87.716174	Banpojung	41.975707	-87.715609	Korean Restaurant
7	Albany Park\n	41.971937	-87.716174	T-Mobile	41.968751	-87.713158	Mobile Phone Shop
8	Albany Park\n	41.971937	-87.716174	Subway	41.968748	-87.712861	Sandwich Place
9	Albany Park\n	41.971937	-87.716174	Dunkin'	41.968255	-87.712964	Donut Shop

Dataframe listing Names, Latitude/Longitude of Neighborhoods, and Venues Information in Chicago.

```
print (Chicago_venues.shape)
```

```
(4451, 7)
```

Information of 4451 Venues in Chicago.

4. METHODOLOGY

- Final Dataframe.
  - Extracting only restaurants from venue category list and create a new dataframe called Chicago_restaurants.

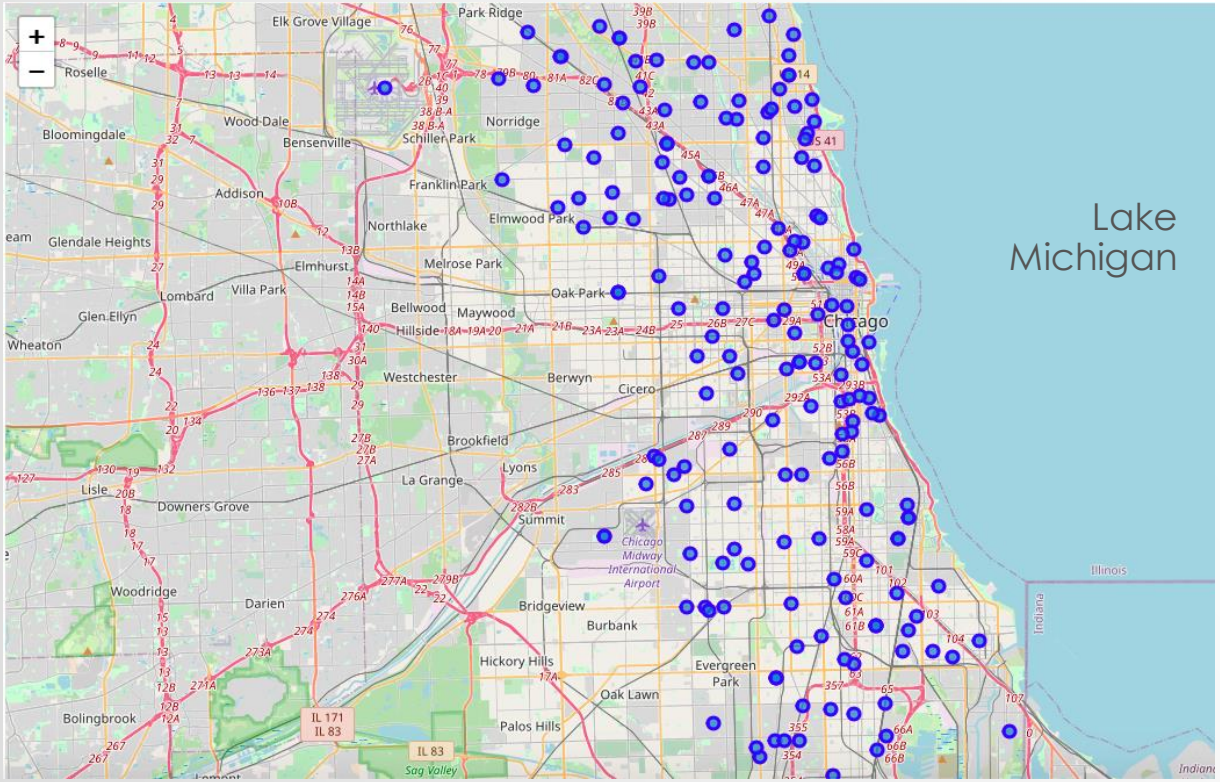
]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
3	Albany Park\n	41.971937	-87.716174	Peking Mandarin Resturant	41.968292	-87.715783	Chinese Restaurant
6	Albany Park\n	41.971937	-87.716174	Banpojung	41.975707	-87.715609	Korean Restaurant
13	Riverdale\n	41.654864	-87.600446	Garden Fast Food	41.656362	-87.597713	Fast Food Restaurant
15	Edgewater\n	41.977139	-87.669273	Taste of Lebanon	41.976151	-87.668847	Middle Eastern Restaurant
23	Edgewater\n	41.977139	-87.669273	Ora	41.975715	-87.668389	Sushi Restaurant
28	Edgewater\n	41.977139	-87.669273	Andies Restaurant	41.977720	-87.668364	Mediterranean Restaurant
31	Edgewater\n	41.977139	-87.669273	Polygon Cafe	41.976375	-87.668551	Asian Restaurant
34	Edgewater\n	41.977139	-87.669273	Calo Ristorante	41.979356	-87.667952	Italian Restaurant
38	Edgewater\n	41.977139	-87.669273	Big Jones	41.979509	-87.668251	Southern / Soul Food Restaurant
39	Edgewater\n	41.977139	-87.669273	Jin Ju	41.976305	-87.668467	Korean Restaurant
43	Edgewater\n	41.977139	-87.669273	Reza's	41.977925	-87.668356	Middle Eastern Restaurant
45	Edgewater\n	41.977139	-87.669273	Vincent	41.979804	-87.667818	New American Restaurant
50	Edgewater\n	41.977139	-87.669273	Octavio Cantina & Kitchen	41.978418	-87.668472	Mexican Restaurant
53	Edgewater\n	41.977139	-87.669273	Anteprima	41.978566	-87.668364	Italian Restaurant
61	Edgewater\n	41.977139	-87.669273	It's Greek To U	41.981288	-87.668106	Greek Restaurant

Dataframe with Name, Latitude/Longitude of Neighborhoods, and ONLY RestauranteVenues Information in Chicago.

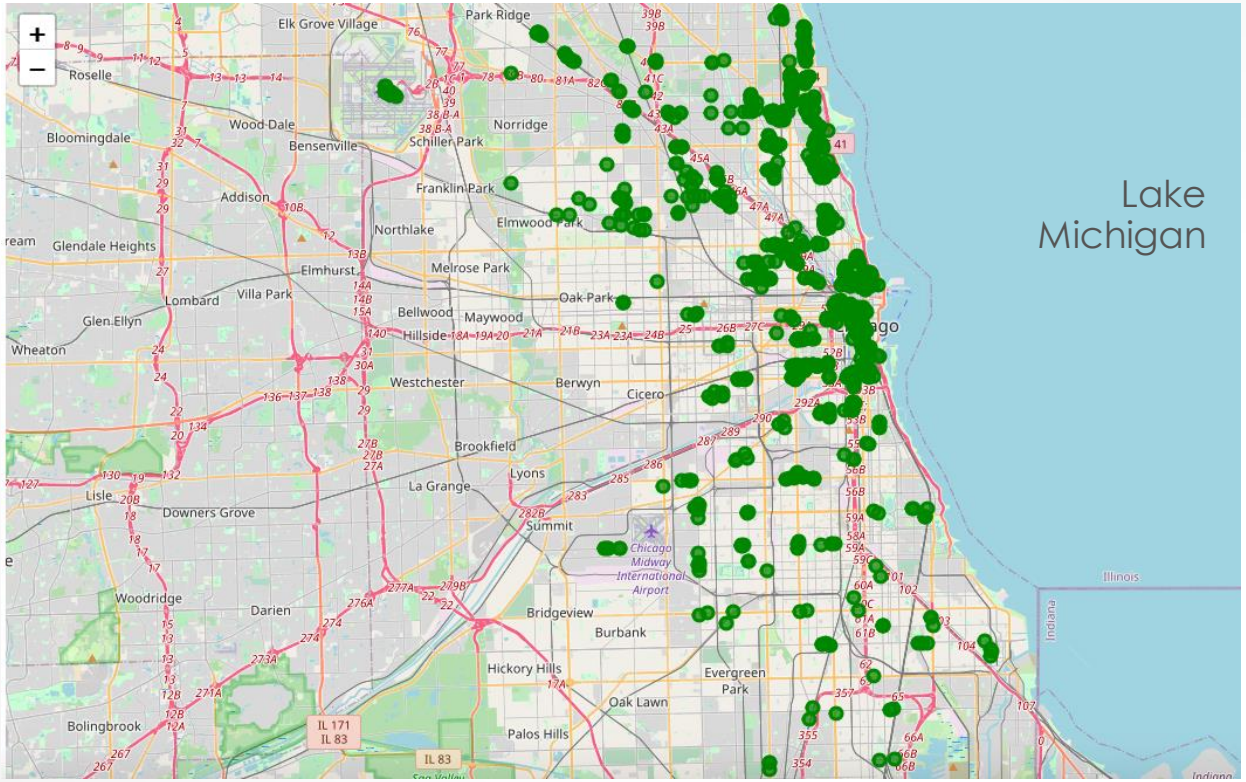
4. METHODOLOGY

Generation of Neighborhood and Restaurants distribution maps in Chicago City is useful to visualize which zones of the city have high and low density of restaurants.



Chicago Neighborhoods Map Distribution.

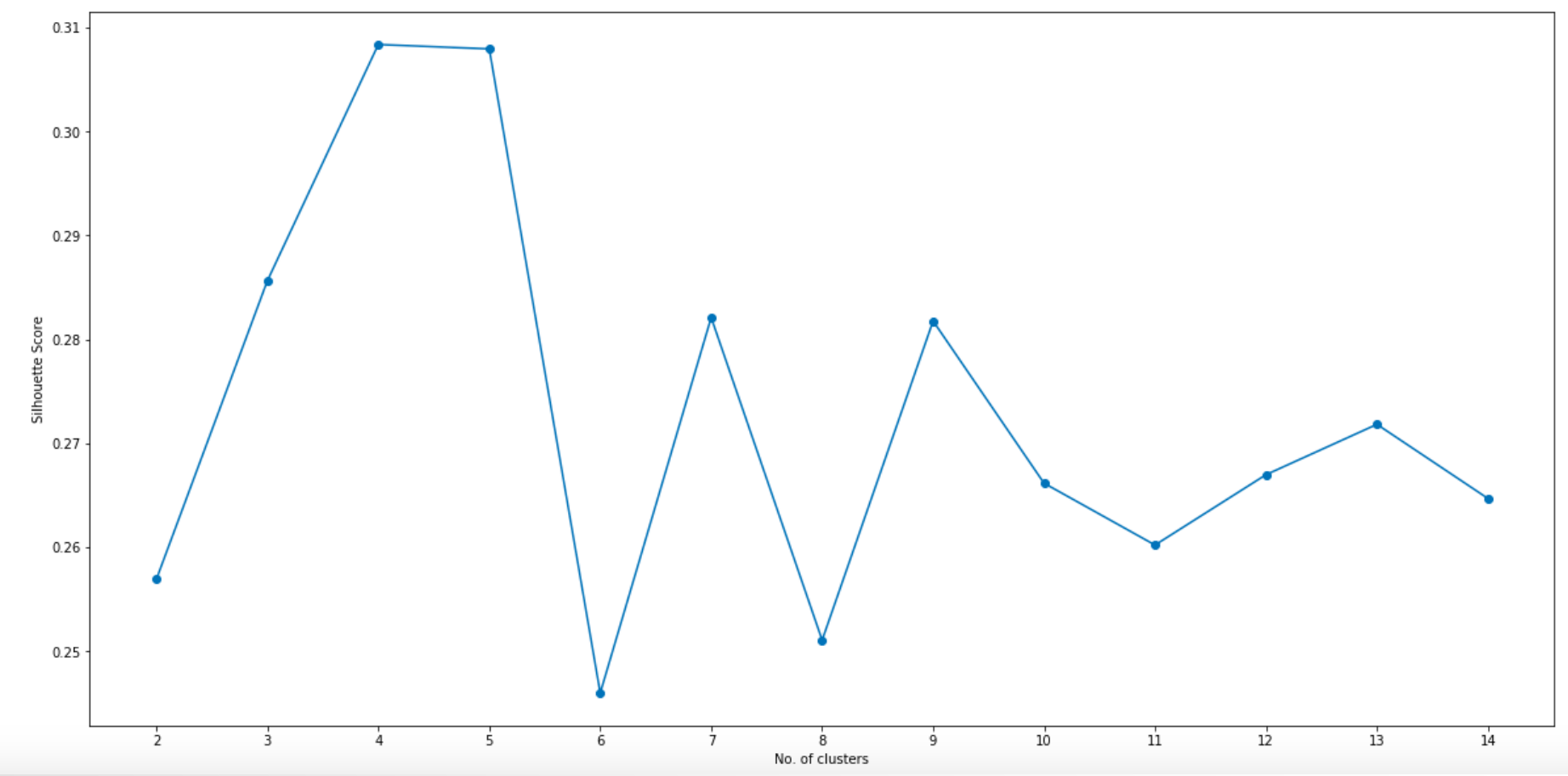
- Neighborhoods distribution.
- Restaurants distribution.



Chicago Restaurants Map Distribution.

4. METHODOLOGY

Applying Machine Learning methodology, an unsupervised classification algorithm (K-Means) for clustering the neighborhoods, taking in account the information obtained from restaurant venues. Previous to apply the algorithm, was calculated the optimal number of cluster, and set in K = 4.



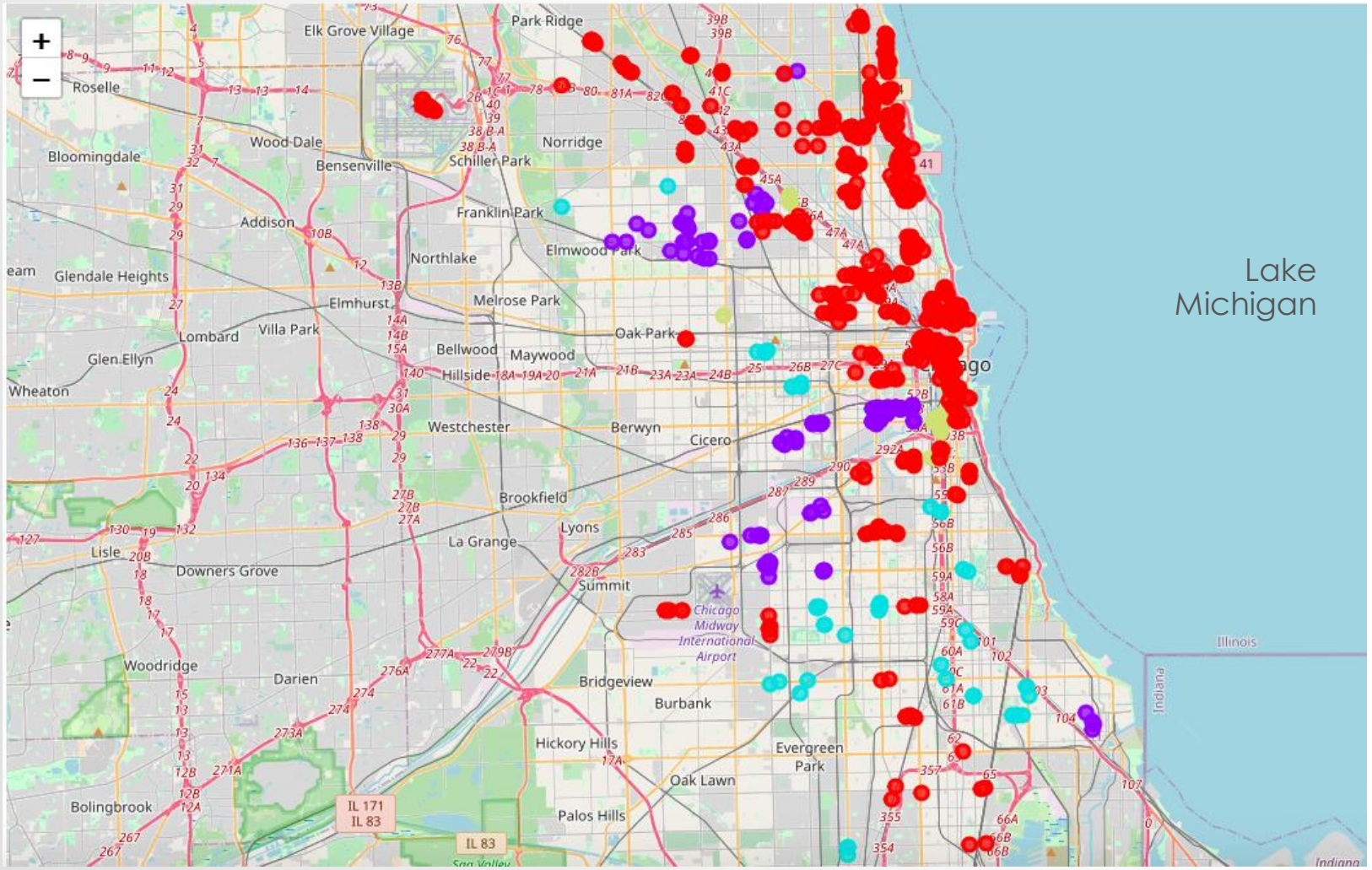
4. METHODOLOGY

Create a dataset which is ordered by the ten most popular restaurant categories by neighborhood in Chicago.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Albany Park\n	Korean Restaurant	Mexican Restaurant	Middle Eastern Restaurant	Cuban Restaurant	American Restaurant	Thai Restaurant	Chinese Restaurant	Seafood Restaurant	English Restaurant	Greek Restaurant
1	Archer Heights\n	Mexican Restaurant	Vietnamese Restaurant	Eastern European Restaurant	Hotpot Restaurant	Halal Restaurant	Greek Restaurant	German Restaurant	French Restaurant	Filipino Restaurant	Fast Food Restaurant
2	Armour Square\n	Chinese Restaurant	Asian Restaurant	Dim Sum Restaurant	Korean Restaurant	Italian Restaurant	Taiwanese Restaurant	Fast Food Restaurant	Hotpot Restaurant	Ramen Restaurant	English Restaurant
3	Ashburn\n	Fast Food Restaurant	Italian Restaurant	Mexican Restaurant	Vietnamese Restaurant	Eastern European Restaurant	Halal Restaurant	Greek Restaurant	German Restaurant	French Restaurant	Filipino Restaurant
4	Auburn Gresham\n	Fast Food Restaurant	American Restaurant	Greek Restaurant	Mexican Restaurant	Chinese Restaurant	Seafood Restaurant	Vietnamese Restaurant	Eastern European Restaurant	Halal Restaurant	German Restaurant
5	Austin	Chinese Restaurant	Vietnamese Restaurant	Eastern European	Hotpot Restaurant	Halal Restaurant	Greek Restaurant	German Restaurant	French Restaurant	Filipino Restaurant	Fast Food Restaurant

4. METHODOLOGY

The cluster classification was performed, having the cluster distribution map based on restaurant venues information for Chicago City.



Chicago clusters distribution map based on restaurant venues.

4. METHODOLOGY

Applying an unsupervised classification algorithm (K-Means) for clustering the neighborhoods, taking in account the information obtained from restaurant venues. Previous to apply the algorithm, was calculated the optimal number of cluster, and set in K = 4.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Mo Comm Ven
3	Albany Park\n	41.971937	-87.716174	Peking Mandarin Resturant	41.968292	-87.715783	Chinese Restaurant	0	Korean Restaurant	Mexican Restaurant	Middle Eastern Restaurant	Cuban Restaurant	American Restaurant	Thai Restaurant	Chinese Restaurant	Seafo Restaura
6	Albany Park\n	41.971937	-87.716174	Banpojung	41.975707	-87.715609	Korean Restaurant	0	Korean Restaurant	Mexican Restaurant	Middle Eastern Restaurant	Cuban Restaurant	American Restaurant	Thai Restaurant	Chinese Restaurant	Seafo Restaura
13	Riverdale\n	41.654864	-87.600446	Garden Fast Food	41.656362	-87.597713	Fast Food Restaurant	2	Fast Food Restaurant	Vietnamese Restaurant	Indian Restaurant	Hotpot Restaurant	Halal Restaurant	Greek Restaurant	German Restaurant	Fren Restaura
15	Edgewater\n	41.977139	-87.669273	Taste of Lebanon	41.976151	-87.668847	Middle Eastern Restaurant	0	Mexican Restaurant	Asian Restaurant	Sushi Restaurant	Italian Restaurant	Thai Restaurant	Restaurant	American Restaurant	Gre Restaura
23	Edgewater\n	41.977139	-87.669273	Ora	41.975715	-87.668389	Sushi Restaurant	0	Mexican Restaurant	Asian Restaurant	Sushi Restaurant	Italian Restaurant	Thai Restaurant	Restaurant	American Restaurant	Gre Restaura
28	Edgewater\n	41.977139	-87.669273	Andies Restaurant	41.977720	-87.668364	Mediterranean Restaurant	0	Mexican Restaurant	Asian Restaurant	Sushi Restaurant	Italian Restaurant	Thai Restaurant	Restaurant	American Restaurant	Gre Restaura
31	Edgewater\n	41.977139	-87.669273	Polygon Cafe	41.976375	-87.668551	Asian Restaurant	0	Mexican Restaurant	Asian Restaurant	Sushi Restaurant	Italian Restaurant	Thai Restaurant	Restaurant	American Restaurant	Gre Restaura
34	Edgewater\n	41.977139	-87.669273	Calo Ristorante	41.979356	-87.667952	Italian Restaurant	0	Mexican Restaurant	Asian Restaurant	Sushi Restaurant	Italian Restaurant	Thai Restaurant	Restaurant	American Restaurant	Gre Restaura

5. RESULTS.

- Examining and analyzing clustering.

➢ Cluster 0.

Generation of a Dataframe with information of restaurants venues classified into Cluster 0.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
3	Albany Park\n	41.971937	-87.716174	Peking Mandarin Resturant	41.968292	-87.715783	Chinese Restaurant	0	Korean Restaurant	Mexican Restaurant	Middle Eastern Restaurant	Cuban Restaurant	American Restaurant	Thai Restaurant	Chinese Restaurant	Seafood Restaurant
6	Albany Park\n	41.971937	-87.716174	Banpojung	41.975707	-87.715609	Korean Restaurant	0	Korean Restaurant	Mexican Restaurant	Middle Eastern Restaurant	Cuban Restaurant	American Restaurant	Thai Restaurant	Chinese Restaurant	Seafood Restaurant
15	Edgewater\n	41.977139	-87.669273	Taste of Lebanon	41.976151	-87.668847	Middle Eastern Restaurant	0	Mexican Restaurant	Asian Restaurant	Sushi Restaurant	Italian Restaurant	Thai Restaurant	Restaurant	American Restaurant	Gree. Restaurant

Dataframe with restaurant venues information classified as Cluster Label 0.

5. RESULTS.

- Examining and analyzing clustering.

➢ Cluster 0.

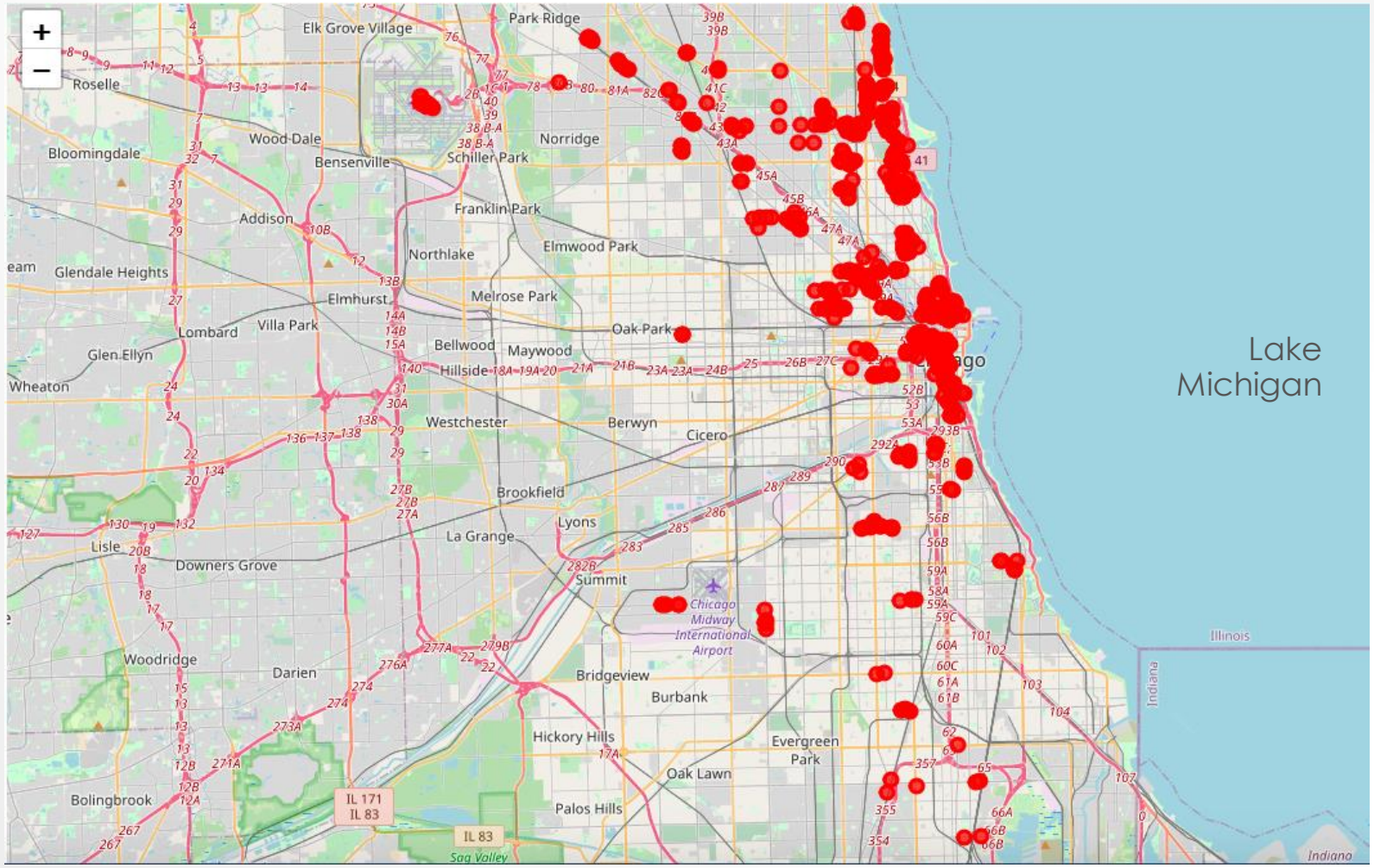
	Most Common Venues
0	American Restaurant
1	Mexican Restaurant

✓ The two **most common** restaurant category obtained from Cluster 0 are classified as American and Mexican Restaurants.

	Less Common Venues
0	Greek Restaurant
1	Restaurant

✓ The two **less common** restaurant category obtained from Cluster 0 are classified as American and Mexican Restaurants.

Chicago Cluster 0 distribution map based on restaurant venues categories.



● Location of Restaurante Venues classified into Cluster 0.

5. RESULTS.

- Cluster 1.

Generation of a Dataframe with information of restaurants venues classified into Cluster 1.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Mos Commor Venue
87	Archer Heights\n	41.811422	-87.726165	El Asador	41.807535	-87.726920	Mexican Restaurant	1	Mexican Restaurant	Vietnamese Restaurant	Eastern European Restaurant	Hotpot Restaurant	Halal Restaurant	Greek Restaurant	German Restaurant	Frenc Restauran
97	Archer Heights\n	41.811422	-87.726165	Taqueria Ochoas	41.807627	-87.727960	Mexican Restaurant	1	Mexican Restaurant	Vietnamese Restaurant	Eastern European Restaurant	Hotpot Restaurant	Halal Restaurant	Greek Restaurant	German Restaurant	Frenc Restauran
199	Belmont Cragin\n	41.931698	-87.768670	Taconazo	41.929640	-87.766087	Mexican Restaurant	1	Mexican Restaurant	Chinese Restaurant	Fast Food Restaurant	Restaurant	American Restaurant	Cuban Restaurant	Vietnamese Restaurant	Englis Restauran
200	Belmont Cragin\n	41.931698	-87.768670	Golden Nugget	41.935193	-87.766229	American Restaurant	1	Mexican Restaurant	Chinese Restaurant	Fast Food Restaurant	Restaurant	American Restaurant	Cuban Restaurant	Vietnamese Restaurant	Englis Restauran
201	Belmont Cragin\n	41.931698	-87.768670	Taqueria Diversey	41.930905	-87.768347	Mexican Restaurant	1	Mexican Restaurant	Chinese Restaurant	Fast Food Restaurant	Restaurant	American Restaurant	Cuban Restaurant	Vietnamese Restaurant	Englis Restauran
207	Belmont Cragin\n	41.931698	-87.768670	La Esquinita	41.931492	-87.769624	Mexican Restaurant	1	Mexican Restaurant	Chinese Restaurant	Fast Food Restaurant	Restaurant	American Restaurant	Cuban Restaurant	Vietnamese Restaurant	Englis Restauran
208	Belmont Cragin\n	41.931698	-87.768670	Yummy Yummy Restaurant	41.931128	-87.767241	Restaurant	1	Mexican Restaurant	Chinese Restaurant	Fast Food Restaurant	Restaurant	American Restaurant	Cuban Restaurant	Vietnamese Restaurant	Englis Restauran
210	Belmont Cragin\n	41.931698	-87.768670	Super Taqueria El Aconaz Restaurant	41.929783	-87.766105	Mexican Restaurant	1	Mexican Restaurant	Chinese Restaurant	Fast Food Restaurant	Restaurant	American Restaurant	Cuban Restaurant	Vietnamese Restaurant	Englis Restauran

Dataframe with restaurant venues information classified as Cluster Label 1.

5. RESULTS.

➢ Cluster 1.

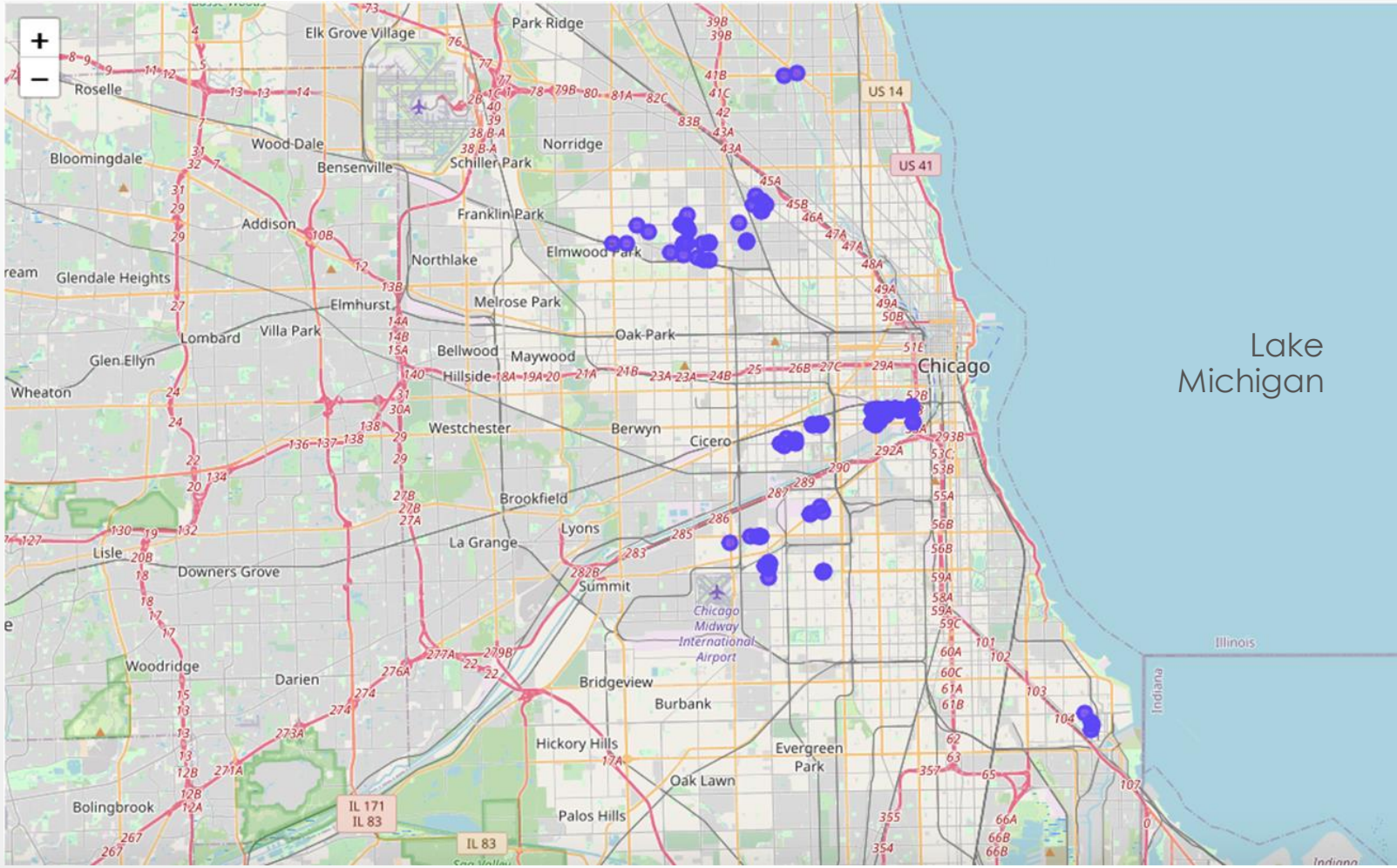
	Most Common Venues
0	Mexican Restaurant
1	Italian Restaurant

- ✓ The two **most common** restaurant category obtained from Cluster 1 are classified as Mexican and Italian Restaurants.

	Less Common Venues
0	German Restaurant
1	Greek Restaurant

- ✓ The two **less common** restaurant category obtained from Cluster 1 are classified as German and Greek Restaurants.

Chicago Cluster 1 distribution map based on restaurant venues categories.



● Location of Restaurante Venues classified into Cluster 1.

5. RESULTS.

- Cluster 2.

Generation of a Dataframe with information of restaurants venues classified into Cluster 2.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Mo Commo Venu
13	Riverdale\n	41.654864	-87.600446	Garden Fast Food	41.656362	-87.597713	Fast Food Restaurant	2	Fast Food Restaurant	Vietnamese Restaurant	Indian Restaurant	Hotpot Restaurant	Halal Restaurant	Greek Restaurant	German Restaurant	Frenc Restaura
111	Ashburn\n	41.749352	-87.713514	Just Tacos Taquera El Sabor de Mexico	41.749818	-87.717886	Mexican Restaurant	2	Fast Food Restaurant	Italian Restaurant	Mexican Restaurant	Vietnamese Restaurant	Eastern European Restaurant	Halal Restaurant	Greek Restaurant	Germa Restaura
124	Avalon Park\n	41.745035	-87.588658	Nita's Gumbo	41.747684	-87.586200	Cajun / Creole Restaurant	2	Fast Food Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Seafood Restaurant	Vietnamese Restaurant	Eastern European Restaurant	Hotpot Restaurant	Hal Restaura
130	Avalon Park\n	41.745035	-87.588658	Just Turkey Restaurant	41.744030	-87.585176	Fast Food Restaurant	2	Fast Food Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Seafood Restaurant	Vietnamese Restaurant	Eastern European Restaurant	Hotpot Restaurant	Hal Restaura
132	Avalon Park\n	41.745035	-87.588658	Mama'z Beef	41.747697	-87.586300	Fast Food Restaurant	2	Fast Food Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Seafood Restaurant	Vietnamese Restaurant	Eastern European Restaurant	Hotpot Restaurant	Hal Restaura
249	Dunning\n	41.937605	-87.833826	Goody Fast Food	41.937588	-87.832983	Fast Food Restaurant	2	American Restaurant	Fast Food Restaurant	Vietnamese Restaurant	Eastern European Restaurant	Hotpot Restaurant	Halal Restaurant	Greek Restaurant	Germa Restaura
291	Morgan Park\n	41.683386	-87.681244	Pepe's Mexican Restaraunt	41.681056	-87.680975	Mexican Restaurant	2	Fast Food Restaurant	Mexican Restaurant	Vietnamese Restaurant	Eastern European Restaurant	Hotpot Restaurant	Halal Restaurant	Greek Restaurant	Germa Restaura
292	Morgan Park\n	41.683386	-87.681244	Joey's Red Hots	41.684206	-87.681230	Fast Food Restaurant	2	Fast Food Restaurant	Mexican Restaurant	Vietnamese Restaurant	Eastern European Restaurant	Hotpot Restaurant	Halal Restaurant	Greek Restaurant	Germa Restaura

Dataframe with restaurant venues information classified as Cluster Label 2.

5. RESULTS.

➢ Cluster 2.

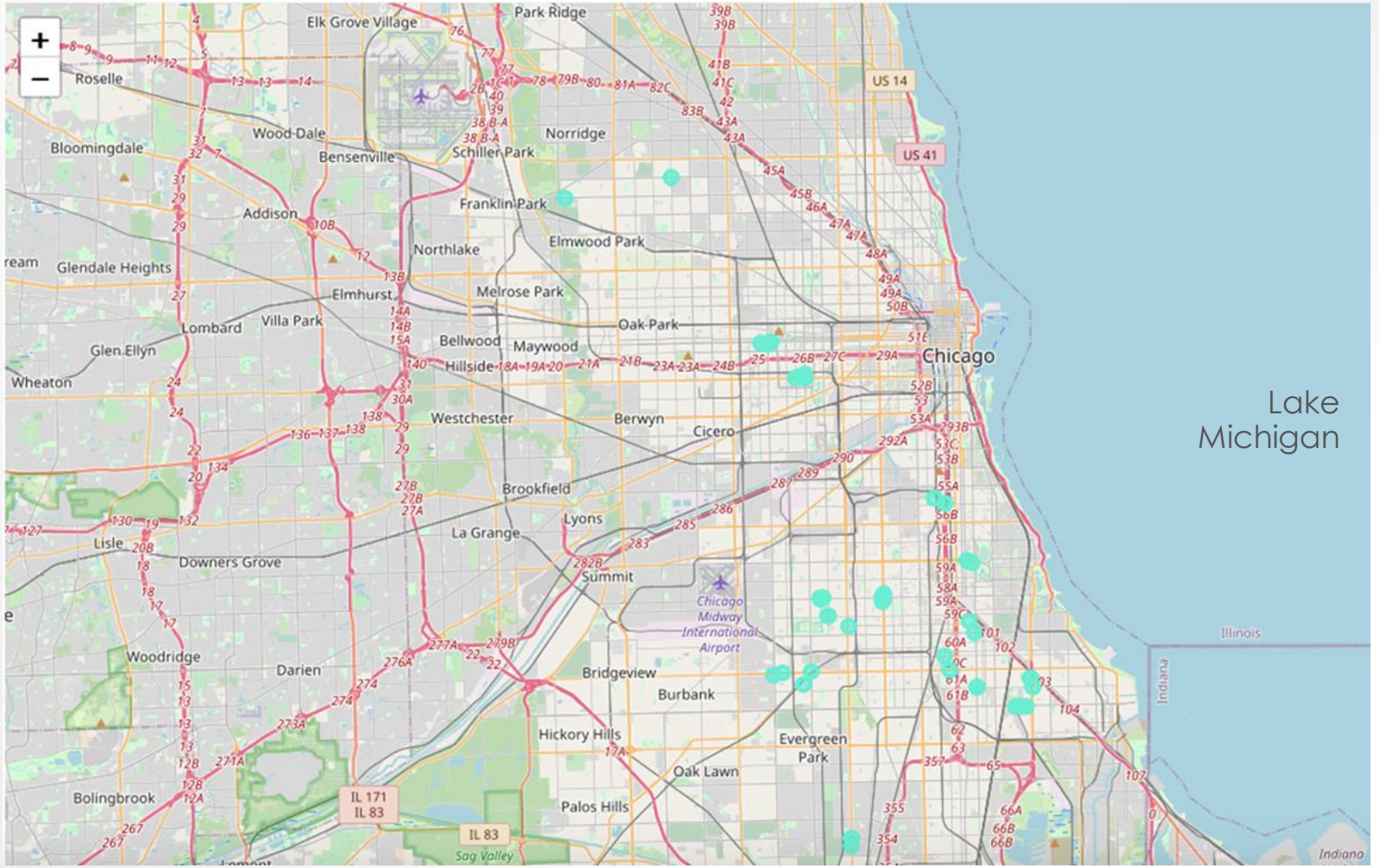
	Most Common Venues
0	Fast Food Restaurant
1	American Restaurant

- ✓ The two **most common** restaurant category obtained from Cluster 2 are classified as Fast Food and American Restaurants.

	Less Common Venues
0	Filipino Restaurant
1	French Restaurant

- ✓ The two **less common** restaurant category obtained from Cluster 2 are classified as Filipino and French Restaurants.

Chicago Cluster 2 distribution map based on restaurant venues categories.



- Location of Restaurante Venues classified into Cluster 2.

5. RESULTS.

- Cluster 3.

Generation of a Dataframe with information of restaurants venues classified into Cluster 3.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
102	Armour Square\n	41.840033	-87.633107	New Furama Restaurant	41.841439	-87.631359	Chinese Restaurant	3	Chinese Restaurant	Asian Restaurant	Dim Sum Restaurant	Korean Restaurant	Italian Restaurant	Taiwanese Restaurant	Fast Food Restaurant	Hotpot Restaurant
103	Armour Square\n	41.840033	-87.633107	Chinese Kitchen	41.838226	-87.637581	Chinese Restaurant	3	Chinese Restaurant	Asian Restaurant	Dim Sum Restaurant	Korean Restaurant	Italian Restaurant	Taiwanese Restaurant	Fast Food Restaurant	Hotpot Restaurant
104	Armour Square\n	41.840033	-87.633107	Franco's Ristorante	41.838273	-87.634155	Italian Restaurant	3	Chinese Restaurant	Asian Restaurant	Dim Sum Restaurant	Korean Restaurant	Italian Restaurant	Taiwanese Restaurant	Fast Food Restaurant	Hotpot Restaurant
105	Armour Square\n	41.840033	-87.633107	Dragon Bowl	41.838410	-87.632549	Asian Restaurant	3	Chinese Restaurant	Asian Restaurant	Dim Sum Restaurant	Korean Restaurant	Italian Restaurant	Taiwanese Restaurant	Fast Food Restaurant	Hotpot Restaurant
107	Armour Square\n	41.840033	-87.633107	Dim Dim Dim-Sum & Bakery	41.841986	-87.631896	Chinese Restaurant	3	Chinese Restaurant	Asian Restaurant	Dim Sum Restaurant	Korean Restaurant	Italian Restaurant	Taiwanese Restaurant	Fast Food Restaurant	Hotpot Restaurant
109	Armour Square\n	41.840033	-87.633107	Ferros	41.838356	-87.632138	Fast Food Restaurant	3	Chinese Restaurant	Asian Restaurant	Dim Sum Restaurant	Korean Restaurant	Italian Restaurant	Taiwanese Restaurant	Fast Food Restaurant	Hotpot Restaurant
110	Armour Square\n	41.840033	-87.633107	Macau Doulao Cruisine	41.841856	-87.632131	Asian Restaurant	3	Chinese Restaurant	Asian Restaurant	Dim Sum Restaurant	Korean Restaurant	Italian Restaurant	Taiwanese Restaurant	Fast Food Restaurant	Hotpot Restaurant
142	Avondale\n	41.938921	-87.711168	중부시장	41.941935	-87.712229	Korean Restaurant	3	Chinese Restaurant	Korean Restaurant	Vietnamese Restaurant	Eastern European Restaurant	Hotpot Restaurant	Halal Restaurant	Greek Restaurant	German Restaurant
148	Avondale\n	41.938921	-87.711168	Eat First #1 Chinese Restaurant	41.939157	-87.711639	Chinese Restaurant	3	Chinese Restaurant	Korean Restaurant	Vietnamese Restaurant	Eastern European Restaurant	Hotpot Restaurant	Halal Restaurant	Greek Restaurant	German Restaurant

Dataframe with restaurant venues information classified as Cluster Label 3.

5. RESULTS.

➢ Cluster 3.

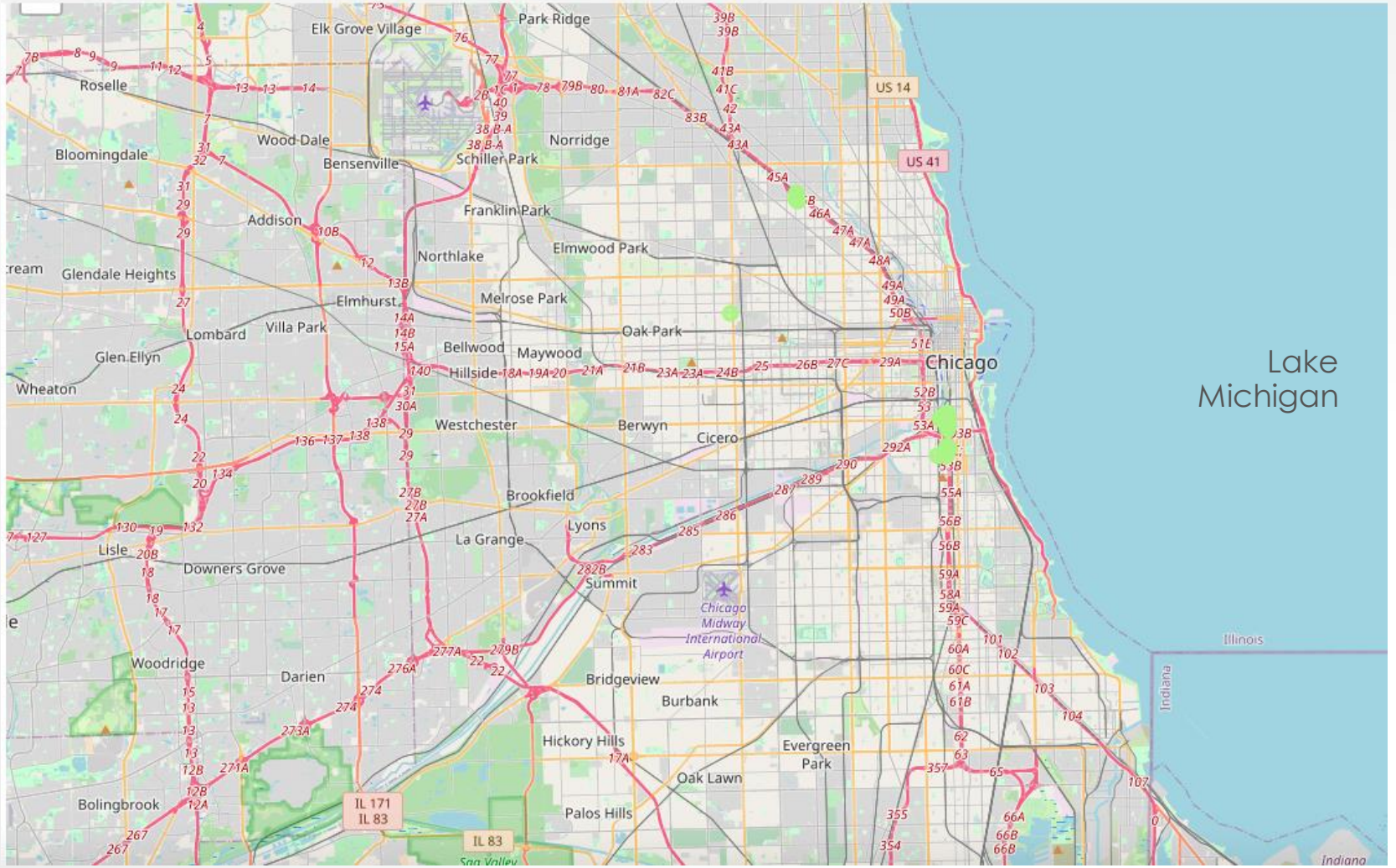
	Most Common Venues
0	Chinese Restaurant

✓ The **most common** restaurant category obtained from Cluster 3 is classified as Chinese Restaurant.

	Less Common Venues
0	English Restaurant
1	Filipino Restaurant

✓ The two **less common** restaurant category obtained from Cluster 3 are classified as English and Filipino Restaurants.

Chicago Cluster 3 distribution map based on restaurant venues categories.



● Location of Restaurante Venues classified into Cluster 3.

6. DISCUSSION.

- Most restaurants in Chicago have been grouped into Cluster 0, where its two most popular restaurant categories are American and Mexican categories, while the less common venue is classified as Greek.

- Cluster 0 is extend around most of the city, mainly in central and northern areas, these areas has the greatest concentrations of business and tourist sights.

- Cluster 1 has Mexican and Italian restaurants as its most common restaurant venues, while the two less common restaurant venues are German and Greek Restaurants. This Cluster is mostly located in two main areas, near to the downtown and very close to the financial district and tourist sighs.

- Cluster 2 is mostly distributed to the southern of Chicago, this area is most known as industrial. The two most common restaurants venues are classified as Fast Food and American. Less popular are Filipino and French restaurants.

- The most popular restaurant category for Cluster 3 is Chinese restaurant, while the two less common categories are English and Filipino.

Cluster Label	Most Common Venue	Less Common Venue
0	American Restaurant	Greek Restaurant
0	Mexican Restaurant	Restaurant
1	Mexican Restaurant	German Restaurant
1	Italian Restaurant	Greek Restaurant
2	Fast Food Restaurant	Filipino Restaurant
2	American Restaurant	French Restaurant
3	Chinese Restaurant	English Restaurant
3	NaN	Filipino Restaurant

Most and Less Common Restaurant Categories obtained from Cluster Classification Calculated Using the K-Means Unsupervised Algorithm.

7. CONCLUSIONS / RECOMENDATIONS .

- Most popular restaurants in Chicago has been classified as Mexican, American and Italian categories. These restaurants are located along most places in the city, likewise these kind of restaurants are located close or into the financial district and tourist sights.

- Open such kind of restaurants in Chicago would be very profitable because are constantly in demand by tourist and locals. it's very remarkable the fact that there are a lot of Mexican origin population in Chicago, also we can see people from all around the world, for this reason would be recommendable to open a new Mexican restaurant in the city if the company offer a great distinguishing factor. The same conclusion would apply to open a new American or Italian restaurant.

- Fast Food Restaurants are the most popular restaurant category found in Cluster 2. This cluster is mainly distributed in Southern and Center Chicago, therefore is very convenient to consider the possibility to open new fast food restaurants to the northern and area.

- Cluster 3 is mainly represented by Chinese restaurants and it is located to the financial district (downtown). Open new Chinese restaurants to other cities neighborhoods would be very profitable.

- Is highly recommended to complement this results and conclusions with other kind of studies, such as marketing and financial feasibility.